利用多时序激光点云数据提取棉花表型参数方法

阳　旭[1,2]，胡松涛[1]，王应华[4]，杨万能[3,4]，翟瑞芳[1*]

（1. 华中农业大学 信息学院，湖北武汉 430070；2. 深圳市财富趋势科技股份有限公司，湖北武汉 430070；3. 华中农业大学 作物遗传改良国家重点实验室，湖北武汉 430070；4. 华中农业大学 植物科学技术学院，湖北武汉 430070）

摘　要：当前，能够实现作物表型参数高效、准确的测量和作物生育期表型参数的动态量化研究是表型研究和育种中亟待解决的问题之一。本研究以棉花为研究对象，采用三维激光扫描LiDAR技术获取棉花植株的多时序点云数据，针对棉花植株主干的几何特性，利用随机抽样一致算法（RANSAC）结合直线模型完成主干提取，并对剩余的点云进行区域增长聚类，实现各叶片的分割；在此基础上，完成植株体积、株高、叶长、叶宽等性状参数的估计。针对多时序棉花激光点云数据，采用匈牙利算法完成相邻时序作物点云数据的对齐、叶片器官对应关系的建立。同时，对各植株表型参数动态变化过程进行了量化。本研究针对3株棉花的4个生长点的点云数据，分别完成了主干提取、叶片分割，以及表型参数测量和动态量化。试验结果表明，本研究所采用的主干提取及叶片分割方法能够实现棉花的枝干和叶片分割。提取的株高、叶长、叶宽等表型参数与人工测量值的决定系数均趋近于1.0；同时，本研究实现了棉花表型参数的动态量化过程，为三维表型技术的实现提供了一种有效的方法。

关键词：棉花表型参数；LiDAR；主干提取；叶片分割；点云数据对齐；三维表型

中图分类号：TP391.4　　**文献标志码：**A　　**文章编号：**202102-SA003

引用格式：阳旭，胡松涛，王应华，杨万能，翟瑞芳．利用多时序激光点云数据提取棉花表型参数方法[J]. 智慧农业(中英文)，2021，3 (1): 51-62.

YANG Xu, HU Songtao, WANG Yinghua, YANG Wanneng, ZHAI Ruifang. Cotton phenotypic trait extraction using multi-temporal laser point clouds[J]. Smart Agriculture, 2021, 3 (1): 51-62. (in Chinese with English abstract)

1　引　言

近年来，随着农作物表型组学和表型技术的发展，三维表型技术由于能够获取比二维表型技术多一个维度的信息，逐渐引起了高通量表型研究人员的关注和兴趣。其中，激光雷达（Light Detection and Ranging，LiDAR）技术作为一种通过发射激光脉冲快速获取目标空间数据的有效手段，既能无损地获取目标对象的三维信息，也克服了传统的计算机视觉技术重建作物的算法复杂度高和耗时长等不利因素，为三维空间中表型参数的自动测量提供了一种新思路。农作物是一种典型的形态结构随着时间变化而不断发生变化的对象，对不同生育期的作物进行持续的点云数据采集，既能够获取作物某一生长点的表型参数，也可以实现作物表型参数的动态变化过程，为育种学家提供更具参考价值的表型数据。

现阶段，作物的三维点云信息主要通过2种

收稿日期：2021-02-01　修订日期：2021-02-26

基金项目：中央高校基本科研业务费资助（2662019PY085）；国家自然科学基金面上项目（31770397）

作者简介：阳　旭（1993—），男，硕士，工程师，研究方向为数据处理。E-mail：yang_9325@163.com。

*通讯作者：翟瑞芳（1979—），女，博士，副教授，研究方向为LiDAR数据处理。电话：027-87288509。E-mail：rfzhai@mail.hzau.edu.cn。

途径获取。一是利用基于图像的计算机视觉方法重建作物目标，如葡萄[1]、番茄[2]、大豆[3]、玉米[4]以及植物根系[5]等。基于图像的计算机视觉方法一般通过相机标定、特征提取、特征匹配以及光束法平差等关键步骤实现，对获取的图像有约束性要求，如保证相邻视点影像间较高的重叠度，且生成的点云数据质量和图像数量关系较大。当作物本身缺乏足够特征时，需引入结构光等获得较丰富的纹理特征，以期获得较高精度的点云数据。Nguyen等[6]搭建结构光系统获取植物多视影像，获得多视点云，并完成点云间的配准，构建了植株的完整三维模型；在此基础上，完成叶片数量、株高、叶片大小和节间距等参数的测量。Ni等[7]利用三台相机获取120幅蓝莓簇图像数据，重建蓝莓点云及模型。针对获取的二维图像数据，利用Mask-RCNN实现单粒蓝莓的分割及成熟度检测，并将分割结果反投回三维空间，在三维空间中完成各蓝莓表型参数的提取。二是利用LiDAR技术直接获取目标三维信息。该类技术首先获取作物对象不同视点的三维点云数据，再利用最临近点迭代法（Iterative Closest Point，ICP）完成多视点的点云配准，生成目标完整的三维模型。利用LiDAR技术开展三维表型研究可概括为2个方面：点云数据处理方法和表型参数提取方法研究。从作物尺度进行划分，LiDAR技术可分别应用在单株作物和大田作物中，对应地获取不同尺度的表型参数。针对大田作物，提取的表型参数主要包括平均冠层高度、投影面积、植株密度、生物量等信息。Jiang等[8]对田间获取的激光点云数据研发了一套和蓝莓收获相关的表型参数获的三维表型技术，包括大小、尺度等。Jin等[9]提出了一种均值归一化向量生长算法（Median Normalized Vector Growth，MNVG）实现田间玉米的茎叶分割，完成了叶面角、茎高、株高等表型参数的测量。Sun等[10]针对重建的田间成熟期棉花植株，利用密度聚类方法实现了棉铃球的分割和计数，为棉花估产提供数据支持。刘守阳等[11]通过构建的数字化植物表型平台（Digital Plant Phenotyping Platform，D3P）获取田间小麦的三维冠层场景，记录了绿色叶面积指数、平均倾角和散射光截获率等信息，在此基础上模拟虚拟LiDAR实验，生成了三维点云数据，展示了LiDAR数据在田间冠层光截获和冠层结构方面的潜力。针对单株作物，主要采用近距离或手持式激光扫描对象，从叶、茎、甚至根等器官角度完成各表型参数的提取。其中，点云配准和点云分割分别是获取三维点云模型和提取表型参数的关键步骤。马福峰等[12]提出了一种基于植株点云邻域几何特征约束的配准方法，选取关键点并估计其中每个点的支撑邻域估算出邻域几何特征，利用几何特征实现点云的初始配准。Chaudhury和Barron[13]将不同视点的点云投影至二维平面上，通过定位节点位置来建立点云间的初始对应关系，该算法实现的前提是点云数据中存在直线特征。点云分割是获取表型参数的前提。适用于作物器官的点云分割方法主要3种：①基于三维区域增长的方法，通过利用区域内部点云的相似属性，制定合适的区域增长规则来完成，主要应用在植株叶片的分割中，如油菜叶片[14]、绿萝和龟背竹[15]等；②基于模型拟合的点云分割方法，利用植株中的直线、圆柱体、平面等典型几何模型来完成，在作物主茎的分割中应用较多，如向日葵[16]和玉米主茎[17]等；③基于机器学习的方法，代表性的算法包括支持向量机（Support Vector Machine，SVM）和均值漂移（Mean Shift）等各种聚类算法。Wahabzada等[18]在获取植株模型的基础上，使用无监督聚类的方法，完成了葡萄、小麦和大麦的叶片和茎秆的分割。喻垚慎等[19]融合了点云数据的多维特征，利用人工标记数据为样本，结合SVM分类器实现了树木枝叶分离，属于有监督分类方法。

按照一定的时间间隔多次采集作物对象的点云数据即可得到作物的多时序点云数据，这些数据中包含了作物对象的时空形态变化信息。针对多时序作物点云数据的研究可为作物表型参数的

动态变化提供强有力的手段和方法，同时，也可为构建作物的生长模型提供数据基础。Su等[20]使用激光扫描仪获取了玉米6个生长阶段的点云数据，计算了玉米株高、叶面积和投影叶面积3个表型参数，并研究表型参数的动态变化趋势。Chaudhury等[21]构建了基于激光扫描仪的近景视觉系统，对野生拟南芥和大麦植株完成不同视点的点云完成配准，通过构建表面不规则网实现植株基于网格的面积计算和体积计算，并绘制了植株从观测第1天到第22天的面积和体积连续变化曲线。Sun等[22]使用激光扫描系统对大田棉花完成了全生育期的表型参数提取和动态变化分析，探讨了表型参数和产量之间的关系。An等[23]对拟南芥进行了连续10天的二维影像数据采集，分别从二维影像和三维模型的角度对比叶片表型参数测量结果的差异性，结果表明与传统的二维影像相比，三维模型能够获得精度更高的表型参数。以上研究表明，虽然三维表型技术已经取得了一定的进展，但针对不同作物植株，数据处理方法各异，因此，其潜力还有待进一步挖掘。

棉花是中国四大经济作物之一，是纺织工业的主要原材料，最大程度地提高棉花产量是栽培和育种的首要目标，对中国的农业经济发展意义重大，而各生育期的表型参数可为该目标的实现提供数据基础。为进一步挖掘三维表型技术在单株棉花植株上的潜力，本研究以单株棉花为研究对象，利用激光扫描技术获取多时序点云数据，研究主干和叶片的分割算法，实现株高、叶长、叶宽、叶面积，以及棉花植株体积等表型参数的计算。在此基础上，实现棉花植株表型参数的动态量化过程，从而为三维空间的高通量表型组学及构建作物生长模型提供方法支撑和数据依据。

2 材料与方法

2.1 点云数据获取

本研究所使用ZGScan手持式激光扫描仪（武汉中观自动化科技有限公司）获取3株棉花共12个时间点的点云数据。在扫描过程中使用的辅助扫描工具包括激光扫描标记点、支架以及扫描数据输入和存储的惠普笔记本电脑。使用戴尔计算机进行后期点云数据分析和处理。该计算配置为：Windows 64位操作系统，16 GB运行内存，处理器为主频3.4 GHz的Inter（R） Core i7-3770 CPU。图1分别为移植后的棉花植株和扫描设备。棉花植株于2018年8月10日在实验室中进行培养，待长至约16 cm后移至栽培盆中种植。



（a）移栽后的棉花植株



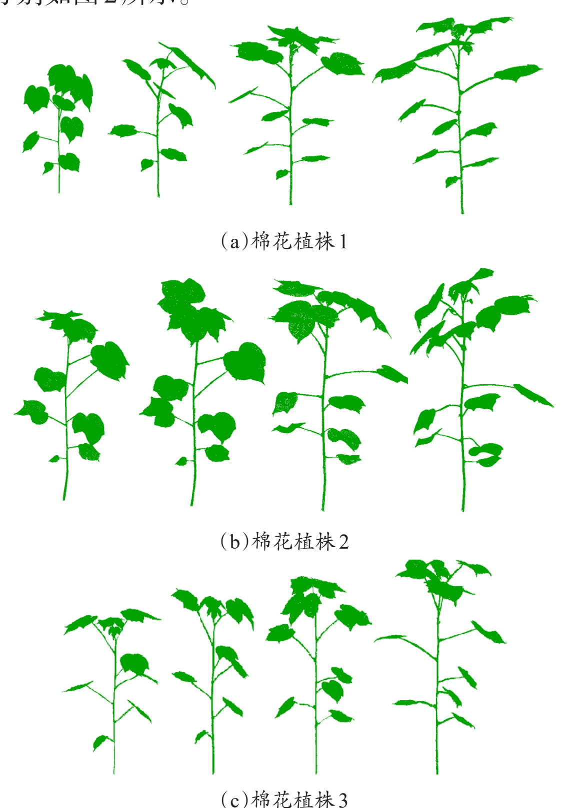
（b）3D激光扫描设备

图1 移栽后的棉花植株和3D激光扫描设备

Fig. 1 The transplanted cotton plant and 3D laser scanner

株高是作物形态学重要参数之一，叶片是作物进行光合作用和呼吸作用的主要载体。棉花生长前期，叶和茎秆的生长最活跃。茎秆的生长，叶片数和叶面积的增长为后期的营养生长和生殖生长积累大量的有机物，是影响棉花高产的重要因素之一。鉴于茎、叶对作物生长发育的重要性，本研究主要探讨LiDAR技术在棉花株高、叶长、叶面积，以及体积等重要表型参数测量方面的潜力，针对3株棉花的4个时间点进行点云数据采集，具体为种植之后的第46天、第49天、第58天和第63天，共采集12组棉花点云数据。随着棉花植株的生长，其形态结构发生了变化，主要表现为植株渐高，部分新叶萌出，部分叶片面积增大，以及少数老叶脱落，叶倾角渐大，节间距有少许增大等趋势。使用Geomagic Studio 2013软件对采集的原始棉花点云数据进行预处

理，包括删除栽培盆、支架等噪声数据。完成棉花植株点云数据预处理后，棉花植株的点云数据分别如图2所示。



(a)棉花植株1

(b)棉花植株2

(c)棉花植株3

注：每横排从左至右分别为第46天、49天、58天和第63天植株

图2 3株棉花的4个生长点的点云数据

Fig. 2 Point clouds data of three cotton plants at four different growth stages

2.2 棉花器官分割方法

后续研究需要通过作物点云数据提取包括株高、叶长、叶宽、以及体积等一系列表型参数，并对其动态变化进行量化。因此，需要对棉花植株点云数据进行器官分割，即将完整的棉花植株点云数据分割成若干互不相交的点云子集。

2.2.1 主干提取

由图2可以看出，棉花植株的主干趋于笔直。因此，根据棉花植株的形态结构以及点云数据特点，采用基于模型的检测方法——随机采样一致性（Random Sample Consensus，RANSAC）算法结合直线模型提取主干部分。RANSAC算法是一种通过随机采样观测数据估计模型参数的方法。首先，在原始数据集中随机指定一组数据集合，设定一个适合该数据的参数模型以及置信度参数。然后，通过迭代选择原始数据的随机子集来不断优化模型的参数。达到迭代终止条件后，选取局内点数量最多的模型即为所求结果。对于多分枝株型的作物，可以采取多次使用RANSAC的方式，直到所有的类直线型枝干都被提取出为止。

2.2.2 叶片分割

在使用RANSAC方法对植株进行主干分割后，得到包括多个叶片的彼此分离的点云数据，只需要将叶片点云进行聚类即可。因此，对提取主干后的剩余点云数据使用欧式聚类分割算法即可得到独立的叶片。欧式聚类分割算法的步骤如下。

（1）为输入的点云数据集P创建K-D树；

（2）构建一个聚类C的空列表以及需要检验的点的队列Q；

（3）在当前点云数据集P中随机选取点P_i，执行下面的步骤。

①将P_i添加到当前队列Q；

②对于每一个属于Q中的点P_i：

在半径$r < d_{th}$的范围内搜索与点P_i相邻的点的集合P_k^i；

③对于每一个属于P_k^i集合中的点，检查当前点是否已经被处理过，如果没有则将其添加到队列Q中；

④当队列Q中的所有点全部处理后，将队列Q中的点添加到聚类C中，同时将队列Q重置为空；

（4）当P中的所有点都被处理完并且属于聚类C中的一部分时结束算法。

2.3 相邻生育期作物点云配准及叶片对应关系建立

为研究同一株作物生长变化及相邻时期表型参数动态变化，需要将相邻时期的作物点云数据进行配准[24]。主要包括两个步骤：（1）点云数

据粗配准；（2）建立作物叶片器官间的一一对应关系。

2.3.1 相邻生育期点云粗配准

点云配准根据精度可分为粗配准和精配准。粗配准是源点云与目标点云在完全未知任何初始相对位置情况下的配准。该操作的主要目的是在初始条件未知的情况下，快速估算一个大致的点云配准矩阵。较典型的是基于全局搜索策略的代表算法采样一致性初始配准（SAmple Consensus Initial Alignment，SAC-IA）算法。该算法在源点云与目标点云之间，随机选取几何特征一致的点组成点对而非所有存在对应关系的组合，通过计算对应点对的变换关系得到最优解[25]。

首先，计算点云的快速点特征直方图（Fast Point Feature Histograms，FPFH）。通过参数化查询点与邻域点之间的空间差异，形成描述查询点的K邻域几何属性的多维直方图。

计算每个点的FPFH后，使用采样一致性初始配准算法完成点云粗配准，具体如下。

（1）从待配准点云S中选取n个采样点，为保证所采样的点具有不同的FPFH特征，采样点两两之间的距离应满足大于预先给定最小距离阈值d；

（2）在目标点云R中查找与点云S中采样点具有相似FPFH特征的一个或多个点，从这些相似点中随机选取一个点作为点云S在目标点云R中的一一对应点；

（3）计算对应点之间的变换矩阵，通过求解对应点变换后的"距离误差和"函数来判断当前配准变换的性能。此处的距离误差和函数多使用Huber误差函数表示，记为$H(l_i)$，见公式（1）。

$$H\left(l_i\right)=\begin{cases}\frac{1}{2}l_i^2, & \left\|l_i\right\|<m_l \\ \frac{1}{2}m_l\left(2\left\|l_i\right\|-ml\right), & \left\|l_i\right\|>m_l\end{cases} \quad (1)$$

其中，m_l为预先给定值；l_i为第i组对应点变换之后的距离差。上述配准的最终目的是在所有变换中找到一组最优解，使得误差函数的值最小，此时的变换即为最终的配准变换矩阵，进一步可得到配准结果。

2.3.2 相邻生育期棉花叶片对应

为量化作物表型参数的动态变化过程，在作物点云数据粗对齐的基础上，需要建立植株叶片间的对应关系。

对于同一株作物T_i和T_{i+1}时期的两组点云，作物T_i时期叶片数为m，T_{i+1}时期叶片数为n。随着植物的生长，可能会发生新叶长出或老叶脱落等现象，因此，叶数m和n可能不同，下面将分别讨论叶片数m和n相同和不同的两种数据处理情况。



注：从棉花点云距离阈值左至右分别是1.0、5.0、6.0和18.0 mm

（a）棉花点云在不同距离阈值下的主干提取结果



注：第从左到右分别为待分割叶片点云最小聚类点数和距离阈值 n_{min}=100，t_d=80 mm；n_{min}=800，t_d=80 mm；n_{min}=100，t_d=120 mm；n_{min}=800，t_d=120 mm

（b）待分割叶片点云和不同最小聚类点数和距离阈值的分割效果

图3 棉花主干提取和叶片聚类结果

Fig. 3 Results of main stem extraction and leaf clustering

（1）当m=n时。计算T_i和T_{i+1}时期植株每个叶片的中心点，然后分别计算T_i时期每个叶片中心点到T_{i+1}时期每个叶片中心点的三维欧氏距离，将T_i和T_{i+1}时期叶片中心点的三维欧式距离构建一个$m\times n$的邻接矩阵$\boldsymbol{M}$，$M(i,j)$表示T_i时期第i个叶片中心到T_{i+1}时期第j个叶片中心的三维欧式距离。设有2组叶片分别为L_1、L_2、L_3和L_1'、L_2'、L_3'，图3描述了这两组叶片各中心点的欧式距离。矩阵的行标和列标分别代表T_i和

T_{i+1}时期的各叶片编号。T_i和T_{i+1}时期植株的点云数据处于同一个坐标系下，使用匈牙利算法（Hungarian Algorithm）来对矩阵$\boldsymbol{M}$求解，使得相邻时期叶片中心点的三维欧式距离之和最小，此时即为相邻时期植株叶片间的最佳匹配，如表1中加粗距离值所示。可得出结论，L_1和L_2'为同一组叶片，L_2和L_1'为同一组叶片，而L_3则和L_3'形成配对关系。

表1 不同时间点同一植株叶片欧式距离矩阵构建

Table 1 Distance matrix among all the leaves of one cotton plant at two different time points

T_i时叶片	T_{i+1}时叶片		
	L_1'	L_2'	L_3'
L_1	3.0	**2.0**	4.0
L_2	**1.0**	2.0	3.0
L_3	2.5	3.0	**2.0**

（2）当$m < n$时。对于同一株作物T_i和T_{i+1}两个时期植株有效叶片数目不同的情况，将T_i时期的叶片数目补充为n个，同时将T_i时期$n-m$个叶片中心到T_{i+1}时期每个叶片中心的距离都记为0，并构建一个$n \times n$的矩阵，然后使用匈牙利算法来对该矩阵进行求解，使得T_i和T_{i+1}两时期叶片中心三维欧式距离之和最小，此时得到T_i时期m个叶片与T_{i+1}时期m个叶片的对应关系。T_{i+1}时期剩余$n-m$个叶片在T_i时期无叶片与其对应，则记为新长出的叶片。

各生育期的棉花叶片间一一对应关系建立后，即可完成株高、叶长、叶宽等一系列表型参数的计算，并可以构建各表型参数的时序变化图。

3 试验和结果分析

3.1 棉花主干和叶片分割结果

RANSAC实现作物主干提取的理想的结果是叶柄与主干分离，主干部分的噪声少并且主干以外的部分叶片数与人工计数结果相同。在使用RANSAC算法实现作物主干提取过程中，不同的参数对结果产生不同的影响。点到模型的距离阈值过小，则提取到的主干不全，距离阈值过大则不能正确地提取主干部分。如果迭代次数过少，则不能得到最优的模型局内点，迭代次数过大会增加算法运行时间。因此，合适的参数选择至关重要。在迭代次数固定的情况下，分别讨论了阈值为1.0、6.0和18.0 mm的主干提取结果，如图3（a）所示。经过多次试验发现当迭代次数为2000次，距离阈值设置在5.0~10.0 mm时，如表2所示，可以得到较为理想的棉花植株主干提取结果。

表2 棉花株高测量结果

Table 2 The experimental results of plant height

结果评定	1号植株				2号植株				3号植株			
时间/天	46	49	58	63	46	49	58	63	46	49	58	63
人工量测/mm	364.60	371.10	460.80	520.10	391.50	426.50	515.60	578.20	312.80	344.80	415.60	480.60
株高计算值/mm	362.50	369.50	459.30	517.90	389.20	424.30	513.40	574.80	310.40	341.60	413.30	477.90
绝对误差/mm	2.10	1.60	1.50	2.20	2.30	2.20	2.20	3.40	2.40	3.20	2.30	2.70
相对误差/%	0.58	0.43	0.33	0.42	0.59	0.52	0.43	0.59	0.77	0.93	0.55	0.56

对于多分枝株型的作物，可以采取多次使用RANSAC的方式，直到所有的类直线型枝干都被提取出为止。使用RANSAC算法进行分割的评价标准根据目的而定。在本试验中，目视评价指标是主干应能反映出作物的株高；量化RANSAC算法提取主干精度的评价指标是主干实际长度与提取主干长度的比值，比值越接近1，精度相对越高。

欧式聚类分割算法的理想结果是分割后的聚类的叶片数量与人工计数结果相同且分割结果中叶片完整。算法实现棉花叶片的分割过程中，需要设置三个参数，近邻搜索半径、最小聚类点数

和最大聚类点数。最小聚类点数和最大聚类点数用于排除噪声以及避免输出近邻搜索半径设置不佳时的错误结果，根据叶子的点云分布情况，一般最大聚类点设置大于点云数量的一半、最小聚类点设置不超过点云数量的10%；近邻搜索半径和点与点之间的距离均值有关，如果近邻搜索半径过大，则点云无法分割，如果过小则会导致分割结果错误。经过一系列的试验验证，各参数如下设置时，可以取得较好的分类结果：最小聚类点数为800个，最大聚类点数为155,000个，近邻搜索半径为120 mm。如图3（b）所示，最小聚类点数为100个时，主干提取后留下的非叶片点云噪声被单独分割出来，对最终结果产生影响；距离阈值为80 mm时，由于叶片区域平均点间距大于叶柄区域，阈值小于叶片区域平均点云间距时，叶片被分割为大量细碎的点，因此在最小聚类点数的限制下只有较为密集的叶柄部分和叶片中的密集区域会被输出。故当提高最小聚类点数到800个后，非叶片点云噪声被排除；而提高距离阈值到120 mm后，叶片与叶柄便可以构成一个整体。

在当前点云数据集中，种子点P_i可随机选取，由于聚类针对的是去除主干后的点云数据，各个叶片已有一定距离间隔，因此，只需要选择合适的距离阈值和迭代次数，聚类后的叶片点集数量均保持不变。

3.2 棉花叶片对应关系建立

利用2.3节介绍的方法分别实现不同生育期作物点云数据及叶片间的一一对应关系。图4是1号棉花第46天和第49天叶片中心三维欧式距离构建的矩阵。经过使用匈牙利算法计算后得到最佳匹配，阴影框内数值表示最佳匹配时对应叶片中心的三维欧式距离，棉花植株T_{i+1}时期9号叶片在T_i时期无叶片与其对应，为新长出的叶片。

多时序的棉花植株间各叶片的对应关系建立结果如图5所示，不同时序点云中的同一叶片采用同一颜色渲染。可以注意到，顶部的新生叶片未能在前期数据中找到对应叶片，而植株底部的红色叶片随着植株的长大成为老叶脱落。

		T_{i+1}								
		3	6	4	8	1	5	7	2	9
T_i	1	121.20	283.44	179.28	328.37	10.63	240.78	298.09	80.88	290.61
	2	162.95	276.51	129.88	292.76	87.94	221.50	241.90	10.68	262.89
	3	28.60	210.56	176.31	280.50	93.95	199.53	276.69	141.38	233.20
	4	178.03	182.35	14.46	209.93	176.58	193.06	136.26	112.72	162.49
	5	147.14	213.58	208.95	154.10	217.02	40.77	213.51	213.70	166.77
	6	161.60	44.15	143.42	189.17	238.02	198.56	181.13	224.01	103.60
	7	241.07	168.59	135.84	91.29	280.75	151.49	39.32	225.57	82.45
	8	271.94	148.41	203.23	58.08	295.70	110.64	146.56	269.94	69.52
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

图4 1号棉花第46天和第49天叶片中心点计算的最佳匹配

Fig. 4 The best match between leaves of No. 1 Cotton Plant on day 46th and day 49th

3.3 表型参数计算及动态变化

3.3.1 株高和体积

通过在提取主干过程中得到的作物主干部分来测量棉花作物的株高。分别获取主干的最低点和最高点的三维坐标，两者之间的欧氏距离即为植株高度。将通过点云计算的作物表型参数与人工使用卷尺测量的表型参数进行对比，从而评价通过作物点云数据提取作物表型参数的准确性。精度评估通过绝对误差、相对误差及线性回归决定系数R^2（Coefficient of Determination）来实现。其中绝对误差可描述为作物点云数据计算的表型参数值与人工测量参数值的差值；相对误差可描述为作物表型参数的绝对误差与人工测量的参数值之比乘以100%，如公式（2）所示。

$$E_R = |Y - L|/L \times 100\% \tag{2}$$

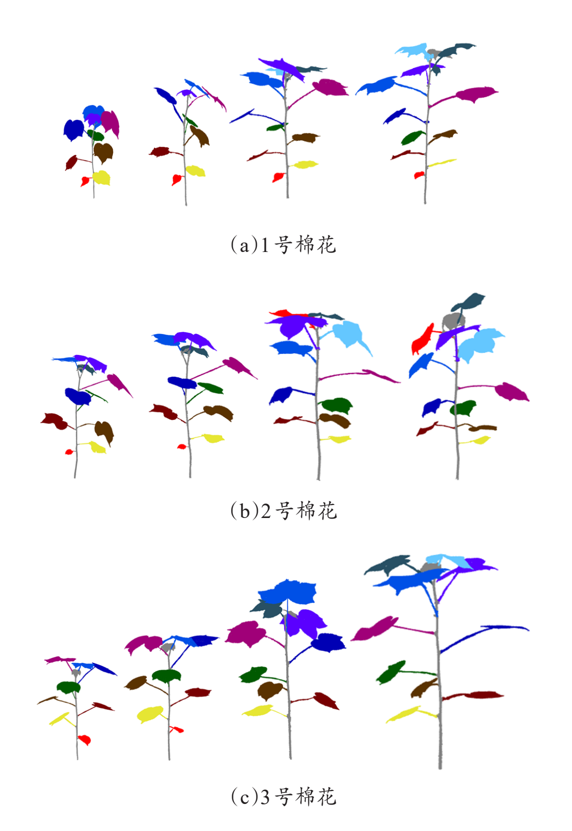
其中，E_R为相对误差；Y为点云数据计算的表型参数值；L为人工测量参数值。

R^2用于度量作物点云数据计算的表型参数与人工测量参数的相关性，公式（3）所示。

$$R^2 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2 - \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \tag{3}$$

其中，n为作物植株叶片数；Y_i为人工测量的表型参数值；$\bar{Y}$为人工测量表型参数值的平均值；$\hat{Y}_i$为点云计算表型参数值。

包含给定三维空间中点集的最小凸多面体被定义为这些点的凸包，凸包的生成和三维空间中



(a)1号棉花

(b)2号棉花

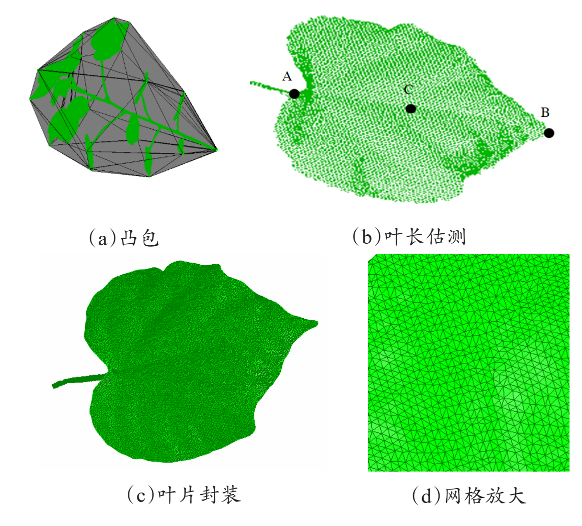
(c)3号棉花

注：从左至右分别为叶片对应后的作物点云在第46天、49天、58天、63天的结果图示。同一叶片用同一种颜色渲染

图5 3株棉花4个时间点各叶片间的对应关系

Fig. 5 Leaf correspondence of three cotton plants at four different growth stages

点云的分布紧密相关。因此，最小包围凸包常用来表示三维空间中作物植株的体积。棉花体积计算过程：输入原始植株点云数据，通过MATLAB获得棉花植株的凸包，即可自动计算凸包的体积，如图6（a）所示。



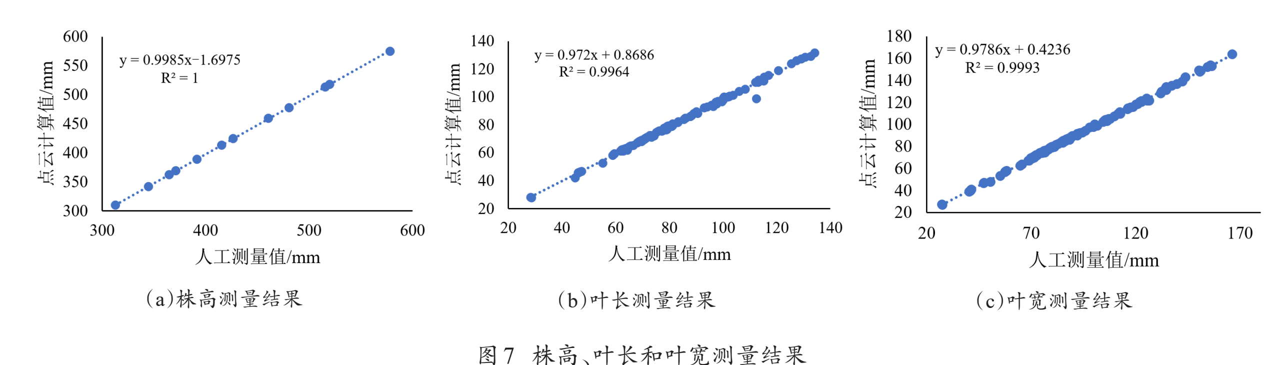
(a)凸包 (b)叶长估测

(c)叶片封装 (d)网格放大

图6 棉花体积、叶长和叶面积估测

Fig. 6 Estimation for cotton plant volume, leaf length and leaf area

图7为对株高、叶长、叶宽等参数进行精度评定，通过对12株棉花所有计算出的株高（如表2所示），叶长和叶宽与人工量测数据进行比较（图7）。可以看出，这三者与人工测量值具备了高度的相关性，株高、叶长和叶宽的决定系数分别为1.0000、0.9964、0.9993，均趋近于1.0。该结果表明通过三维激光扫描技术可获取作物的高精度表型参数。主要由于：1）所使用的近景激光扫描系统扫描精度较高，可达0.03 mm；2）与基于影像的二维表型量测方法相比，三维激光点云数据多了一个维度的信息，因而在表型参数测量上更具优势。



(a)株高测量结果 (b)叶长测量结果 (c)叶宽测量结果

图7 株高、叶长和叶宽测量结果

Fig. 7 Measurment results of plant height, leaf length and leaf width

3.3.2 叶长、叶宽和叶面积估算

棉花叶片存在一定弧度的弯曲。因此，直接利用叶片两端点的欧式距离值来表描述叶长并不适用。本研究采用一种逐渐逼近的方法来测量叶片长度[26]。假设叶片两端点分别为A（X_A，Y_A，Z_A）和B（X_B，Y_B，Z_B），AB两端点间的距离为L_{AB}，取AB中点C（图6（b）），分别计算点A、C和点B、C之间的距离l_{AC}和l_{BC}。令$l_{AB}=l_{AC}+l_{BC}$，设定l_{AB}和L_{AB}间长度偏差的阈值d，比较l_{AB}和L_{AB}；若$\left|l_{AB}-L_{AB}\right|<d$，令：

$$l_{AB}=L_{AB}+d \tag{4}$$

则l_{AB}为叶长；否则分别对AC和BC继续处理。用B替换AC段中端点标号C，AC段更新为AB段。该过程一直迭代执行，直到距离差值满足阈值条件为止。叶宽也采用逐渐逼近法完成计算。

对棉花植株点云数据进行叶片器官分割得到每个独立的叶片点云数据，使用Geomagic Studio 2013软件对叶片点云数据进行封装，获得叶片的网格数据，之后计算各网格的面积之和，即为叶片的面积。在网格化过程中，需要对网格数量和距离值进行约束，如果网格数量越多，则目标表面越精细，计算出的面积值与真实值越接近。本研究设置的网格数量为2,500,000个，距离值为0，因此，能够保证得到足够多的网格数量，也即较高精度的面积估值。封装后的叶片如图6（c）和6（d）所示。

3.3.3 表型参数动态变化估计

棉花植株点云数据获取过程中进行了多次的作物点云数据采集，因此可以将同一作物不同生育期提取的表型参数按照先后顺序进行对比分析，从而实现作物相邻时期表型参数变化过程，并将其结果可视化，为育种和遗传学相关研究提供数据基础。

3株棉花的4个时间序列的株高、体积的变化趋势如图8（a）和8（b）所示，叶长、叶宽、叶面积的动态变换趋势分别如图8（c）～8（k）所示。从图中可以看出，该观测期内，株高的变化趋势较为平缓，呈稳步上升态势。体积的增长趋势和株高较相似，主要原因在于体积的计算和株高相关性较强。老叶面积变化不大，新叶面积增加趋势明显，如1号棉花中顶部的三片叶子（图5（a））中分别渲染为紫色、深蓝色、蓝绿色。叶面积和叶长的变化趋势较为一致，同样因为叶面积与叶长有较强的相关性。

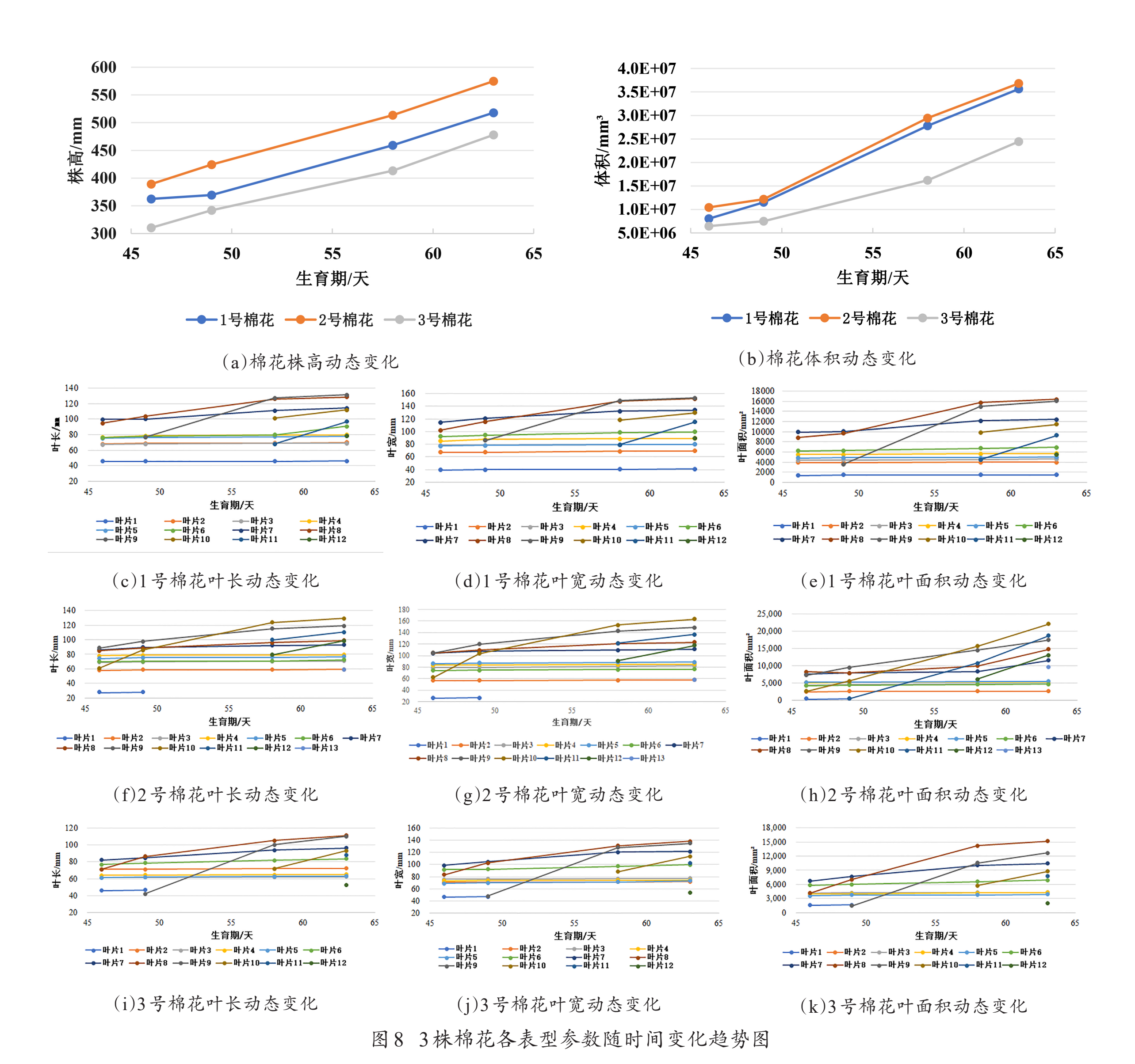
本研究中，从1株棉花植株的点云数据采集开始到输出各表型参数结果值，累计耗时约30 min，其中数据采集约15～20 min，主干提取和叶片聚类在30 s内完成，体积计算约11 s，单叶的叶面积估算约耗时30 s。此过程中，数据采集较为耗时，该过程与采集对象的复杂度有关，对象越复杂，则耗时越长。

4 结论与讨论

为挖掘三维激光扫描技术在作物表型组学中的应用，本研究利用激光扫描技术获取了棉花的多时序三维激光点云数据，实现了三维空间内株高、叶长、叶宽、叶面积及体积等表型参数获取和动态量化。主要取得的结果如下。

（1）利用RANSAC算法结合直线模型提取了3株棉花共12组点云的主干提取，并对剩余点云使用欧式聚类算法完成每一株棉花的叶片分割。完成了棉花植株的叶长、叶宽、叶面积、株高和体积等参数的计算，对比试验结果表明，提取的株高、叶长、叶宽与人工测量值均具有高度的相关性，决定系数R^2趋近于1.0，克服了传统测量方法主观性强、具有一定的破坏性等缺点，为后续产量等研究奠定基础。

（2）获取同一株棉花作物不同生育期的多时序点云数据，研究了作物点云的粗配准和叶片器官的一一对应方法，实现了对应关系的可视化。并根据器官间的对应关系实现了棉花表型参数变化的动态量化过程。通过研究茎秆、叶片不同时期的动态变化，可为剖析棉花个体的器官生长发育过程提供精确的信息，也可为构建作物生长的动态模型提供可靠的数据基础。



(a)棉花株高动态变化

(b)棉花体积动态变化

(c)1号棉花叶长动态变化

(d)1号棉花叶宽动态变化

(e)1号棉花叶面积动态变化

(f)2号棉花叶长动态变化

(g)2号棉花叶宽动态变化

(h)2号棉花叶面积动态变化

(i)3号棉花叶长动态变化

(j)3号棉花叶宽动态变化

(k)3号棉花叶面积动态变化

图8 3株棉花各表型参数随时间变化趋势图

Fig. 8 Dynamic change of phenotypic traits of the three cotons

本研究的不足之处在于，由于侧重于方法研究，采集数据时，暂未覆盖棉花植株的完整生育期，无法开展棉花植株的全生育期的各表型参数的获取和动态变化研究。因此，在后续工作中，将主要利用全生育期的棉花点云数据，实现全生育期表型参数的动态变化可视化和深度数据挖掘。同时，本研究方法也可推广至包含直线型的作物对象如玉米、油菜等的表型研究中。

参考文献：

[1] BRIGLIA N, WILLIAMS K, WU D, et al. Image-based assessment of drought response in grapevines[J]. Frontiers in Plant Science, 2020, 11: 595.

[2] ROSE J C, PAULUS S, KUHLMANN H. Accuracy analysis of a multi-view stereo approach for phenotyping of tomato plants at the organ level[J]. Sensors, 2015, 15: 9651-9665.

[3] GUAN H, LIU M, MA X, et al. Three-dimensional reconstruction of soybean canopies using multisource imaging for phenotyping analysis[J]. Remote Sensing, 2018, 10(8): ID 1206.

[4] WANG Y, WEN W, WU S, et al. Maize plant phenotyping: Comparing 3D laser scanning, multi-view stereo reconstruction, and 3D digitizing estimates[J]. Remote Sensing, 2018, 11(1): ID 63.

[5] MASUDA T. 3D shape reconstruction of plant roots in

a cylindrical tank from multiview images[C]// 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Piscataway, New York, USA: IEEE, 2019: 2149-2157.

[6] NGUYEN T, SLAUGHTER D, MAX N, et al. Structured light-based 3D reconstruction system for plants[J]. Sensors, 2015, 15(8). 18587-18612.

[7] NI X, LI C, JIANG H, et al. Three-dimensional photogrammetry with deep learning instance segmentation to extract berry fruit harvestability traits[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 171: 297-309.

[8] JIANG Y, LI C, TAKEDA F, et al. 3D point cloud data to quantitatively characterize size and shape of shrub crops[J]. Horticulture Research, 2019, 6(1): 43-59.

[9] JIN S, SU Y, WU F, et al. Stem-leaf segmentation and phenotypic trait extraction of individual maize using terrestrial LiDAR data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 57(3): 1336-1346.

[10] SUN S, LI C, CHEE P W, et al. Three-dimensional photogrammetric mapping of cotton bolls in situ based on point cloud segmentation and clustering[J]. Journal of Photogrammetry and Remote Sensing, 2020, 160: 195-207.

[11] 刘守阳, 金时超, 郭庆华, 等. 基于数字化植物表型平台(D3P)的田间小麦冠层光截获算法开发[J]. 智慧农业(中英文), 2020, 2(1): 87-98.

LIU S, JIN S, GUO Q, et al. An algorithm for estimating field wheat canopy light interception based on digital plant phenotyping platform[J]. Smart Agriculture, 2020, 2(1): 87-98.

[12] 马福峰, 耿楠, 张志毅. 基于邻域几何特征约束的植株三维形态配准方法研究[J]. 计算机应用与软件, 2016, 33(9): 184-189.

MA F, GENG N, ZHANG Z. On 3D plant morphology registration method based on geometrical feature constraint of neighborhood[J]. Computer Applications and Software, 2016, 33(9): 184-189.

[13] CHAUDHURY A B M, BARRON J L. Junction-based correspondence estimation of plant point cloud data using subgraph matching[J]. IEEE Geoscience & Remote Sensing Letters, 2016, 13(8): 1119-1123.

[14] YANG L, ZHAI R, YANG X, et al. Segmentation of plant organs point clouds through super voxel-based region growing methodology[J]. Computer Engineering and Applications, 2019, 55(16): 197-203..

[15] LI D, CAO Y, TANG X, et al. Leaf segmentation on dense plant point clouds with facet region growing[J]. Sensors, 2018, 18(11): ID 3625.

[16] GELARD W, DEVY M, HERBULOT A, et al. Model-based segmentation of 3D point clouds for phenotyping sunflower plants[C]// 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications Porto. Portugal, 2017: ID hal-02733980.

[17] 梁秀英, 周风燃, 陈欢, 等. 基于运动恢复结构的玉米植株三维重建与性状提取[J]. 农业机械学报, 2020, 51(6): 209-219.

LIANG X, ZHOU F, CHEN H, et al. Three-dimensional maize plants reconstruction and traits extraction based on structure from motion[J]. Transactions of the CSAM, 2020, 51(6): 209-219.

[18] WAHABZADA M, PAULUS S, KERSTING K, et al. Automated interpretation of 3D laserscanned point clouds for plant organ segmentation[J]. BMC Bioinformatics, 2015, 16(1): ID 248.

[19] 喻垚慎, 云挺, 杨绪兵. 基于激光点云数据的植物器官多维特征分割方法[J]. 数据采集与处理, 2015, 30(5): 1054-1061.

YU Y, YUN T, YANG X. Multi dimension feature segmentation method of foliage organs based on laser point cloud data[J]. 2015, 30(5): 1054-1061.

[20] SU Y, WU F, AO Z, et al. Evaluating maize phenotype dynamics under drought stress using terrestrial LiDAR[J]. Plant Methods, 2019, 15(1): ID 11.

[21] AYAN CHAUDHURY, CHRISTOPHER WARD, ALI TALASAZ. Machine vision system for 3D plant phenotyping[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2019, 16(6): 2009-2022.

[22] SUN S, LI C, PATERSON A H, et al. In-field high throughput phenotyping and cotton plant growth analysis using LiDAR[J]. Frontiers in Plant Science, 2018, 9: 1-17.

[23] AN N, WELCH S, MARKELZ R, et al. Quantifying time-series of leaf morphology using 2D and 3D photogrammetry methods for high-throughput plant phenotyping[J]. Computers & Electronics in Agriculture, 2017, 135: 222-232.

[24] 阳旭. 基于多时相点云数据的作物表型参数获取及动态量化方法研究[D]. 湖北: 华中农业大学, 2019.

YANG X. Research on crop phenotypic parameters acquisition and dynamic quantification method based on multi-temporal point cloud data[D]. Hubei: Huazhong Agricultural University, 2019.
[25] RUSU R B, HOLZBACH A, BLODOW N, et al. Fast geometric point labeling using conditional random fields[C]// 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, New York, USA: IEEE, 2009: 7-12.
[26] 史蒲娟, 翟瑞芳, 常婷婷, 等. 基于单目视觉和激光扫描技术的油菜植株模型重建及株型参数测量[J]. 华中农业大学学报, 2017, 36(3): 63-68.

Cotton Phenotypic Trait Extraction Using Multi-Temporal Laser Point Clouds

YANG Xu[1,2], HU Songtao[1], WANG Yinghua[4], YANG Wanneng[3,4], ZHAI Ruifang[1*]

（1. *College of Informatics*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*; 2. *Shenzhen Fortune Trend Technology Co.*, *Ltd.*, *Wuhan* 430070, *China*; 3. *National Key Laboratory of Crop Genetic Improvement*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*; 4. *College of Plant Science & Technology*, *Huazhong Agricultural University*, *Wuhan* 430070, *China*）

Abstract: To cope with the challenges posed by the rapid growth of world population and global environmental changes, scholars should employ genetic and phenotypic analyses to breed crop varieties with improved responses to limited resource environments and soil conditions to increase crop yield and quality. Therefore, the efficient, accurate, and non-destructive measurement of crop phenotypic traits, and the dynamic quantification of phenotypic traits are urgently needed for crop phenotypic research, and breeding as well as for modern agricultural development. In this study, cotton plants were taken as research objects, and the multi-temporal point cloud data of cotton plants were collected by using three-dimensional laser scanning technology. The multi-temporal point clouds of three cotton plants at four time points were collected. First, RANSAC algorithm was implemented for main stem extraction on the original point cloud data of cotton plants, then region growing based clustering was carried out for leaf segmentation. Plant height was estimated by calculating the end points of the segmented main stem. Leaf length and width measurements were conducted on the segmented leaf parts. In addition, the volume was also estimated through the convex hull of the original point cloud of plant cotton. Then, multi-temporal point clouds of plants were registered, and organ correspondence was constructed with the Hungarian method. Finally, dynamic quantification of phenotypic traits including plant volume, plant height, leaf length, leaf width, and leaf area were calculated and analyzed. The overall performance of the approaches achieved a matching rate through a series of experiments, and the traits extracted by using of point cloud showed high correlation with the manually measured ones. The relative error between plant height and manual measurement results did not exceed 1.0%. The estimated leaf length and width on point clouds were highly correlated with the manually measured ones, and the coefficient of determination was nearly 1.0. The proposed 3D phenotyping methodology can be introduced and used to other crops for phenotyping.

Key words: cotton phenotypic traits; LiDAR; stem extraction; leaf segmentation; point cloud registration; 3D phenotyping